STAT 300 Chapter 2

Zepu Zhang
January 26, 2012

1 Probability: sample space and events

Probability studies uncertainties and randomness. Examples of actions with uncertain outcome:

1. Toss coins.
2. Throw dice.
3. Draw cards.
4. Measure the air temperature.
5. Pick a student and obtain his/her blood type.

Experiment: an action or procedure whose outcome is uncertain.

Sample space, $\mathcal{S}$, of an experiment: the set of all possible outcomes of the experiment.

Example List the sample space of
(1) toss a coin once;
(2) toss a coin 3 times;
(3) toss 3 identical coins simultaneously;
(4) throw a die twice;
(5) draw 2 cards from a deck of 52 cards and get their types (i.e. heart, club, diamond, spade);
(6) take a measure of the temperature. (The temperature at any given moment is assumed to be an unknown constant. The randomness is due to measurement.)

From these examples we see that an outcome may be a scalar (if numerical), a vector, or a set.

When the outcome contains multiple components, it is important to know or specify whether their order matters. If order matters, the outcome is a vector, e.g. (2) and (4); if order does not matter, the outcome is a set (or un-ordered list), e.g. (3) and (5).

Event: any collection of outcomes in the sample space, that is, a set of outcomes.

simple event, compound event, null event

An event occurs if the outcome of an experiment is in the collection of outcomes that define the event. Therefore, one experiment may trigger multiple events to occur, because that particular outcome may be in definition of multiple events.

Example

In the example (5), define event A as "there is at least one heart in the two cards drawn". Then A contains the following outcomes:

heart-heart, heart-diamond, heart-club, heart-spade

Define event B as "there is at least one club in the two cards drawn". Then B contains

club-club, club-diamond, club-heart, club-spade

If an experiment results in a heart and a club, then both events A and B have occurred.

2 Relations between events

Since an event is a set of outcomes, then relations between events are nothing but relations between sets.

2.1 Set theory

Sample space $\mathcal{S}$: the whole set, whose elements are all possible "outcomes" of the experiment.

Event A: any subset of $\mathcal{S}$, including $\emptyset$ and $\mathcal{S}$.

Complement: $A' = \mathcal{S} \setminus A$.

Union: $A \cup B$.

Intersection: $A \cap B$.

A and B are **mutually exclusive** or **disjoint** if $A \cap B = \emptyset$.

The concepts of union, intersection, and disjoint generalize to multiple (finite or countably infinite) events.

2.2 Venn diagram

Example Fig. 2.1 at the end of Chap. 2.1. Use Venn diagram to illustrate union, intersection, complement, and mutually exclusiveness.

3 Axioms of probability

We want to assign to each event (note: not "outcome") A a number $P(A)$, called "probability", which measures the "chance" that A will occur in an experiment. Such a measure should have the following intuitively reasonable properties. These are not "theorems"; rather, they are "axioms". They are not derived from other things, but rather stipulated from the outset as something to be accepted.

Axiom 1. $P(A) \geq 0$.

Axiom 2. $P(\mathcal{S}) = 1$.

Axiom 3. If <u>infinitely many</u> events A_1,A_2,..., are <u>disjoint</u>, then

$$P(A_1 \cup A_2 \cup \cdots) = \sum_i P(A_i)$$

Proposition 1. $P(\emptyset) = 0$.

Proposition 2. The property in axiom 3 is valid for a <u>finite</u> number of disjoint events.

Example Prove the two propositions above. [Hint: use the axioms!]

Interpretation

The axioms specify the properties that the probability assignment should satisfy. However, for any $\mathcal{S}$, there may be more than one way to assign probabilities to events such that these axioms are satisfied. Therefore one needs to choose an "interpretation" of the probability $P(A)$;

the interpretation will guide one to pick a particular assignment of probabilities to events.

Example Toss a coin. Two possible outcomes: H and T. $S = \{H, T\}$. All possible events: $\emptyset$, $\{H\}$, $\{T\}$, $\{H, T\}$. We already know, by the axioms and propositions, that $P(\emptyset) = 0$ and $P(\{H, T\}) = 1$. How should we determine $P(\{H\})$ and $P(\{T\})$?

> There is no definite answer to this question. These probabilities are not determined, but rather are assigned. As long as the assignment does not violate the axioms, it is valid; however, whether it is reasonable depends on what interpretation we give to the term "probability".

One of the common and simple interpretations is based on relative frequency in repeatable experiments.

Relative frequency and limiting relative frequency: Conduct the experiment N times and let n be the number of times that event A has occurred. The ratio n/N is called the "relative frequency" of the occurrence of event A. Experience indicates that as $N \to \infty$, the relative frequency will stabilize, approaching a limiting relative frequency. Then take this limiting relative frequency as $P(A)$.

Example Re-visit the coin-tossing example. Suppose the coin is balanced. If we toss it many, many, many times, what "relative frequency" the outcome H should have? Intuition tells us it should be 0.5. So, $P(\{H\}) = 0.5$.

Alternatively, intuition tells us that $P(\{H\}) = P(\{T\})$. Then since

$$P(\{H\}) + P(\{T\}) = P(S) = 1$$

we have $P(\{H\}) = P(\{T\}) = 0.5$.

Note 1. This interpretation is meaningful for repeatable experiments and less so for other kinds of uncertainties.

2. Even in repeatable experiments, in reality one can not conduct the experiment infinitely and observe the limiting relative frequency.

3. This interpretation of probability is called "objective" (by some). There are other interpretations, including "subjective" ones, which will not bother us in this course.

The "limiting relative frequency" interpretation of probability is what we use in this course to determine probabilities. This will soon be applied in the situation of "equally-likely" outcomes.

4 Properties of probability

Prove the following properties using the axioms and propositions:

1. $P(A) + P(A') = 1$.
2. $P(A) \le 1$.
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

   Example: Ex. 2.14 in Chap. 2.2.
4. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

   This property can be generalized to more events.

Note 1. A useful relation between any two sets A and B:

$$A = (A \cap B) \cup (A \cap B')$$

In words, A is the union of two disjoint sets, one being the part of A that is also in B and the other being the part of A that is outside of B. (Either of the two disjoint sets may be empty.)

2. Property 4 can be derived from property 3:

$$P(A \cup B \cup C) = P(A \cup B) + P(C) - P\big((A \cup B) \cap C\big)$$

then, use

$$P(A \cup B) = \cdots$$
$$P\big((A \cup B) \cap C\big) = P\big((A \cap C) \cup (B \cap C)\big) = \cdots$$

5 Determining probabilities

5.1 Equally likely outcomes

If an experiment has N possible outcomes and all the outcomes are <u>equally likely</u>, then $P(A) = N(A)/N$, where $N(A)$ is the number of outcomes in event A.

Example Draw one card from a fully mixed deck of cards. There are 52 possible outcomes and it is reasonable to assume all individual cards have the same chance of being picked.

Event A: the card drawn is a spade.

Answer: $N(A) = 13$, hence $P(A) = 13/52 = 1/4$.

5.2 Product (or multiplication) rule

Suppose an outcome consists of an ordered list of k items selected, and there are n_1 choices for item 1, n_2 choices for item 2,..., n_k choices for item k, then there are $n_1 \cdot n_2 \cdots n_k$ possibilities for this outcome.

We will use the rules 5.1 and 5.2 A LOT.

The "ordered list of items selected" often arises in the form of "a sequence of steps" leading to an outcome.

Example Throw 4 dice into 4 spots. How many possible outcomes?

Example Draw 3 cards from one deck and arrange them in the order they are drawn. How many possible outcomes?

Example Experiment: draw 4 cards.
Event: all 4 cards are of the same suit, and their ranks (face number) are in sequence, like 2, 3, 4, 5.
Question: how many possible outcomes does this event contain?

Answer. Think of it as 2 steps: (1) choose the suit—4 choices; (2) choose the ranks—11 choices (1,2,3,4; 2,3,4,5;...;11,12,13,1). $4 \times 11 = 44$.

Or think of it this way: each outcome is uniquely identified by an ordered pair of items: `(suit, starting rank)`.

5.3 The simple property $P(A) = 1 - P(A')$ is very useful

Example Ex. 2.13 in Chap. 2.2.

5.4 Use Venn diagram

5.5 Divide and conquer

Divide the event into disjoint sub-events, determine the probability of each sub-event, then add up.

6 Counting techniques

In an equally-likely-outcome experiment, calculating probabilities amounts to counting the number of outcomes in the desired event, in addition to counting the number of outcomes in the sample space.

Permutation

Taking k out of n items and placing them in order, the number of possible outcomes is

$$P_{k,n} = n \cdot (n-1) \cdot (n-2) \cdots [n-(k-1)] = \frac{n!}{(n-k)!}$$

(Product of k numbers counting down from n.)

Complete permutation: there are $n!$ possible orderings of n items.

Note It appears the notation $P_{k,n}$ is not universal. We follow the textbook.

Combination

Taking k out of n items, the number of unique combinations (ignoring their order) is

$$\binom{n}{k} = \frac{P_{k,n}}{k!} = \frac{n!}{k!(n-k)!}$$

(Reasoning: to turn a permutation to a combination, ignore the difference between the $k!$ orderings of any unique set of k items.)

Example A quality control engineer is to randomly pick 5 out of the 11 available products for inspection. Suppose 2 of the 11 products have defects. What is the probability that (1) no defective product is picked? (2) exactly one defective product is picked?

Example Consider a regular deck of 52 cards. For a five-card poker hand, find the probability of (a) All of different ranks; (b) One pair; (c) Two pairs; (d) Three of a kind: three cards of the same rank and two others of different ranks, for example JJJ74; (e) A straight: five cards in sequence; the ace can be either high or low; (f) A flush: five cards of the same suit; (g) All are spades; (h) There are exactly 2 clubs and 2 hearts.

Exercise Ex. 2.22 in Chap. 2.3.

Exercise Ex. 2.23 in Chap. 2.3.

7 Conditional probability

Example One guy commutes 30 miles to work and makes it on time 60% of the time on average. What is the prob of his getting to work on time on a particular day given that (1) It is snowing? (2) He got up extra early?

Example Pick one American at random, the prob that he or she skies is 0.3%. What if this person is from Fairbanks?

Example Roll a die, $A = \{1\}$. $B = \{1, 3, 5\}$. (1) $P(A)$? (2) $P(A \mid B)$?

Answer: $P(A) = 1/6$. $P(A \mid B) = 1/3$.

Why?

In the case of equally-likely outcomes,

$$P(A \mid B) = \frac{N(\text{in } A\text{, besides being in } B)}{N(\text{in } B)} = \frac{N(A \cap B)}{N(B)}$$

Since B is a condition, that is, it is assumed that B occurs, then the effective sample space for the subsequent (conditional) event is B. Under this condition, the event A is actually $A \cap B$. Hence $P(A \mid B) = N(A \cap B)/N(B)$.

More generally,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

(This can be considered a definition of conditional probability.)

Understand this using Venn diagram.

Example Ex. 2.26 in Chap. 2.4.

Note We often use the definition this way:

$$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$$

Example Draw two cards from a deck of 52 cards. What's the probability of getting a pair?

Answer

Solution 1. $N(S) = \binom{52}{2}$. $N(A) = \binom{13}{1}\binom{4}{2}$.

$$P(A) = \frac{N(A)}{N(S)} = \frac{13 \times 4 \times 3}{1 \times 2} \Big/ \frac{52 \times 51}{1 \times 2} = 1/17$$

Solution 2.

$$A = (1,1) \cup (2,2) \cup \cdots \cup (13,13)$$

These sub-events are mutually exclusive, and they all have the same probability.

$$P(A) = \sum_{i=1}^{13} P(i,i) = 13 \times P(\text{first draw } 1)P(\text{second draw } 1 \mid \text{first draw } 1) = 13 \times \frac{4}{52}\frac{3}{51} = \frac{1}{17}$$

Note

$$P(A) = P(A \cap B) + P(A \cap B') = P(B)P(A \mid B) + P(B')P(A \mid B')$$

This is a special case of the law of total probability (p. 72, 7th ed; p. 78, 8th ed).

Think how this can be used to solve problems by "classification" or "divide and conquer".

Example Pick 2 cards without replacement from a 52 deck.
A1 = an ace is selected on the first draw.
A2 = an ace is selected on the second draw.
Find $P(A_1)$, $P(A_2)$.

Answer

$$P(A_1) = \frac{4}{52}$$

$$P(A_2) = P(A_1)P(A_2 \mid A_1) + P(A_1')P(A_2 \mid A_1') = \frac{4}{52}\frac{3}{51} + \frac{48}{52}\frac{4}{51} = \frac{4}{52}$$

Note: $A_2 = A_1$!
Implication: in a poker game, seating does not matter.

Example Ex. 2.27 in Chap. 2.4.

Example The probability that a randomly selected family belongs to the AAA auto club is 0.25 (*Source*: American Automobile Association). If a family belongs to AAA, the probability that they have more than one car is 0.45. Suppose a family is randomly selected. What is the probability they have more than one car and belong to AAA?

Answer: $A =$ belong to AAA. $M =$ have more than one car.

$$P(M \cap A) = P(A)P(M \mid A) = 0.25 \times 0.45 = 0.1125.$$

Note: It is also true that $P(M \cap A) = P(M)P(A \mid M)$, but it is not useful for this question, because $P(M)$ and $P(A \mid M)$ are unknown.

Example **A Traveling Salesperson** During frequent trips to a certain city a traveling salesperson stays at hotel A 50% of the time, at hotel B 30% of the time, and at hotel C 20% of the time. When checking in, there is some problems with the reservation 3% of the time at hotel A, 6% of the time at hotel B, and 10% of the time at hotel C. Suppose the salesperson travels to this city.

(a) Find the probability that the salesperson stays at hotel A and has a problem with the reservation.
(b) Find the probability that the salesperson has a problem with the reservation.
(c) Suppose the salesperson has a problem with the reservation, what is the probability that the salesperson is staying at hotel A?

8 Independence

If $P(A \mid B) = P(A)$, that is, the probability of A's occurrence is not affected by whether B occurs or not (or, knowing that B occurs gives us no info about the probability of A), we say A and B are independent.

Definition of independence:

$$P(A \cap B) = P(A)P(B)$$

This definition includes the (uninteresting) situations $P(A) = 0$ and $P(B) = 0$.

Note 1. If $P(A) \neq 0$ and $P(B) \neq 0$, the following three relations are equivalent:

$$P(A \cap B) = P(A)P(B)$$

$$P(A \mid B) = P(A)$$
$$P(B \mid A) = P(B)$$

Exercise Prove their equivalence, that is, from one you can derive the other two.

2. Independence is not the same as disjoint. Actually if $A \cap B = \emptyset$, the two events are <u>dependent</u>. Knowing that B has occurred certainly tells us about the occurrence of A: A will not occur.

3. There is no useful way to indicate "independence" by a Venn diagram.

4. In the definition of independence, we can replace any event by its complement. For example, if A and B are independent, then A' and B are independent, because

$$\begin{aligned} & P(A' \cap B) = P(A')P(B) \\ \Leftrightarrow \quad & P(B) - P(A \cap B) = \big(1 - P(A)\big)P(B) \\ \Leftrightarrow \quad & P(A \cap B) = P(A)P(B) \end{aligned}$$

5. Generalization: independence between more than two events (Chap. 2.5). Not required.